Design, Development and Implementation of a Bioinformatics Course Sequence

Gina Cooper
DeVRY University Columbus Campus, 1350 Alum Creek Dr, Columbus, OH 43209
gcooper@devrycols.edu

Abstract - Bioinformatics is emerging as a new and exciting field. Biological data repositories are complex and increasing in size constantly. Manipulating, analyzing, and interpreting this data requires the use of computers and trained biomedical informaticians. Spending on bioinformatics projects has exceeded a billion dollars in the past year and is expected to grow every year. To support this growing need for trained biomedical informaticians, Devry University has developed a bioinformatics curriculum available to both onsite and online students. This paper provides a context for two of the classes in the curriculum, the course content, labs, and instruction.

I. INTRODUCTION

Since the completion of the Human Genome project in 2001, several new information technology projects have surfaced. Due to the open source nature of the bioinformatics community, much of the bioinformatics data is available online. This data is constantly being downloaded, manipulated, and analyzed. In addition to data, numerous bioinformatics tools and algorithms are being developed to assist users in efficiently analyzing bioinformatics data. The need for skilled biomedical informaticians is increasing. A search on the Medzilla search site reveals numerous matches for biomedical informaticians who are competent using tools to analyze biological data [1]. The bioinformatics curriculum at Devry University will prepare graduates to effectively retrieve and analyze biological data. Four courses in the bioinformatics curriculum provide students with the necessary tools to be successful as biomedical informaticians. These courses are flexible enough to be taught in online, onsite, and hybrid online-onsite formats. Two of these courses will be discussed in this paper. These are BMI310 Bioinformatics I with Lab and BMI315 Bioinformatics II with Lab. A third course, BMI325 will follow this sequence and will focus on advanced programming concepts for bioinformatics. Finally, BMI433 represents a capstone course in the curriculum in which students create and document a semester long project for a business.

II. ONLINE BIOINFORMATICS CURRICULUM

Current delivery formats for education can be mixed between online, onsite, and hybrid courses. Students may choose to take some courses online while taking other courses onsite. To maintain the freedom to mix delivery formats, bioinformatics courses must be flexible in both topics and material covered.

The field of bioinformatics is especially suited to the online teaching environment. Most data is freely available for download and manipulation. Tools for sequence alignment, phylogenetics, visualization, and microarray analysis are open source or freely available online. Table 1 lists a few of the many bioinformatics databases publicly available [2]:

<table>
<thead>
<tr>
<th>Subject</th>
<th>Source</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><a href="http://srs.ebi.ac.uk">http://srs.ebi.ac.uk</a></td>
</tr>
<tr>
<td>Protein Structure</td>
<td>Protein Data Bank</td>
<td><a href="http://www.rcsb.org/pdb">http://www.rcsb.org/pdb</a></td>
</tr>
<tr>
<td>Entrez Structure DB</td>
<td>PROWL</td>
<td><a href="http://prowl.rockefeller.edu">http://prowl.rockefeller.edu</a></td>
</tr>
<tr>
<td>Peptide mass spectroscopy</td>
<td>RESID</td>
<td><a href="http://www-">http://www-</a> nibrf.georgetown.edu/pirww/search/textresid.html</td>
</tr>
</tbody>
</table>

TABLE 1

Biological Data and Information Sources
Adapted from Gibas and Jambeck Table 6-1

In addition to the plethora of material available online, there exists a helpful online community where students can benefit from the experience of experts in the field. Molecular biology simulations can also be performed remotely for laboratory experience.

Even with the availability of online databases and tools, challenges exist to the development and implementation of BMI310 and BMI315 in an online delivery format. Following is the discussion of the implementation of these two courses online. The lectures, assignments, labs, discussions and exams enhance learning and provide students with hands-on knowledge to complete their understanding of the topics of the courses.

III. DEVELOPMENT

A. Course Content

The course content of BMI310 Bioinformatics I and BMI315 Bioinformatics II reflect current research and industry trends. Main topics in BMI310 are broad, yet provide enough depth that the students can learn from hands-on projects. Some topics for BMI310 include: the basics of molecular biology, databases and information retrieval, database searches and pairwise alignment, BLAST, FASTA, CLUSTALW, substitution patterns, phylogeny, prokaryotic and eukaryotic genomes, microarrays, protein structure and function, visualization tools such as CN3D, and proteomics and drug design. These topics provide a broad introduction to the field of bioinformatics and impart a good knowledge baseline to the students.

The course content for BMI315 provides more in-depth knowledge in the areas introduced in BMI310 as well as an introduction to Perl programming. This course builds upon the knowledge gained in BMI310. The initial design of BMI315 was a tools class covering online tools such as BLAST, FASTA, CLUSTALW, and CN3D. However, the scope of BMI315 was expanded as students entering BMI325 lacked sufficient programming knowledge in Perl to create efficient and effective programs. Some topics which were added to BMI315 include: molecular biology basics, data structures, Perl programming, UNIX and LINUX scripting, bioPerl, dynamic programming, database homology searches, genetic algorithms, and neural networks. Due to the additional programming concepts added to BMI315, the online tools topics were moved to BMI310. This design works well as the algorithms for these tools are introduced in BMI310.

This course sequence will provide students with basic and advanced knowledge in bioinformatics. Labs and exercises based on the content of these courses are assigned to encourage learning.

In an onsite classroom setting, these topics would be introduced in a lecture format with demonstrations to reinforce the material to students. The online setting provides an opportunity for the lecture and demonstrations to be combined. Lectures are short narratives using images, links, and flash interactives to convey new information and reinforce it with mini-demonstrations.

B. Textbooks and Content Outlines

Many bioinformatics textbooks are designed for the graduate level classes and therefore contain too many advanced topics to be used as undergraduate text books. Ten textbooks were evaluated for BMI310 with evaluation criteria of: topics covered in the book, sample Perl code, clarity for undergraduate students, homework problems, answers to homework problems, and instructor resources. No textbook contained all topics for BMI310, however one textbook: *Fundamental Concepts of Bioinformatics* by Krane and Raymer was chosen based on the criteria above [3]. The nine remaining textbooks were listed as recommended textbooks. Descriptions of each book and their analysis also follow including the text and recommended texts chosen for the course.

### Text:


This text is written in clear, concise, and understandable language for readers in all disciplines, especially undergraduate students. End of chapter problems increase student’s problem solving skills and encourage critical thinking. Answers to odd problems are available in the appendix. The appendix also contains a wide variety of Perl code and examples to aid the students and instructor with example problems. Instructor resources include answers to even-numbered end of chapter questions. However, instructor resources did not contain PowerPoint slides.

### Recommended texts:

A. Lesk. *Introduction to Bioinformatics*. Oxford, 2002. (ISBN 0-19-925196-7). This text is well written and well organized. It provides end of chapter problems and webems. The webems are questions that can be used for extra study and research, however no odd-numbered answers are provided in the back of the book. This